BEBR

FACULTY WORKING
PAPER NO. 1519

Downward Transfers, Completely Monotone Utility and K-th Degree Stochastic Dominance

Donald V. Coes

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 1519

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

December 1988

Downward Transfers, Completely Monotone Utility and K-th Degree Stochastic Dominance

Donald V. Coes, Associate Professor
Department of Economics

I would like to thank H. Markowitz and M. H. Simonsen for helpful discussions of several points which led to this paper.

Downward Transfers, Completely Monotone Utility
and K-th Degree Stochastic Dominance

The classical Pratt-Arrow measure of absolute risk aversion is strengthened here by identifying greater risk with a downward transfer of the mass of the distribution, rather than with variance. Downward transfers may be expressed as sequences of differences in central moments. Utilities for which the risk premium is non-decreasing for all such transfers are completely monotone. This result may then be used to derive stochastic dominance rules up to any degree.

Downward Transfers, Completely Monotone Utility,

and K-th Degree Stochastic Dominance

The measure of absolute risk aversion $r(x) = -U^{(2)}(x)/U^{(1)}(x)$, discovered independently by Pratt [1964] and Arrow [1965], has played a central role in the development of economic theory under uncertainty for more than two decades. The measure, and the widely accepted hypothesis that it is non-increasing in x, are not sufficient in themselves, however, to provide unambiguous answers to a number of issues involving economic behavior under uncertainty. This has led to several extensions of the concept to make it applicable to a wider class of problems.[1]

Pratt and Zeckhauser [1987] have recently proposed a strengthening of the concept to "proper risk aversion", which requires that an undesirable prospect should not be made desirable by the addition of an independent undesirable prospect. Their results show a close, but not equivalent relation between proper risk aversion and "completely monotone" utilities, which are concave utilities whose higher derivatives alternate in sign. Our purpose in this paper is to propose an alternative strengthening of absolute risk aversion which yields completely monotone utilities. In addition to their theoretical

1 Among them are the extension by Kihlstrom and Mirman [1981] to multiple arguments in the utility function, and a strengthening of the concept by Ross [1981] to allow base wealth x to be stochastic. A further extension by Machina [1982] allows the increment to wealth to be stochastic. The Ross strengthening and extensions are discussed in Machina and Neilson [1988].

interest, these utilities are of operational significance, since they permit the derivation of stochastic dominance (SD) rules up to any desired degree, thus expanding considerably the set of uncertain alternatives which may be ordered under these expected utility criteria.

The original Pratt-Arrow measure, as well as the more recent stronger measures of Ross [1981] and Machina and Neilson [1987] are derived by identifying risk with variance, under the explicit assumption that other characteristics of the relevant distributions are negligible. We derive our results using a concept of risk which identifies greater risk with higher probabilities of "worst-case" outcomes, rather than with variance. Our definition of an increase in risk, which we term a "downward transfer" (DT) is developed in Section I. An intuitive interpretation of this definition is that it corresponds to a transfer of some of the mass of the distribution to the lower tail, increasing the likelihood of the most adverse outcomes. We show that DT's may be expressed as sequences of differences in the central moments. Such sequences of moments appear to have been noted first by Fishburn [1980], but have received relatively little attention from economic theorists.

In Section II the risk premium π corresponding to the DT concept is derived. Necessary and sufficient conditions for π to be non-decreasing for any DT imply complete monotonicity. In Section III the restrictions on utility implied by complete monotonocity are exploited to obtain SD rules up to any arbitrary degree.[2] These rules order a larger set of uncertain alternatives than is possible using a second-degree rule, based on concavity, or a third degree rule, for which non-increasing absolute risk aversion in the original Pratt-Arrow sense is sufficient.

I. K-th Degree Downward Transfers

In the succeeding discussion we shall assume that the points of increase of the relevant distribution functions are restricted to the closed interval $[a,b]$. As $-a$ and b may be chosen to be arbitrarily large, this assumption imposes little cost in terms of economic relevance and simplifies considerably our argument. It is sufficient, moreover, for the existence of the moments of the distributions up to any desired degree.

Consider the transfer of some of the mass of a distribution from a neighborhood of the mean to the left or lower tail. More formally, define a downward transfer of the first degree, or DT(1), of the density $f(x)$ having a mean μ, as a transformation of $f(x)$ into $g(x)$ such that $g(x) = f(x) + H_0(x)$, where $H_0(x) \geq 0$, for $x \leq \mu - \epsilon$, $\epsilon > 0$, and $H_0(x) = 0$ for $x > \mu + \epsilon$, with the first inequality strict for at least one x in $[a, \mu - \epsilon]$. Hence $H_0(x) < 0$ for at least one x in the half-open interval $(\mu - \epsilon, \mu + \epsilon]$.

2 Higher degree stochastic dominance relations have been discussed in Fishburn [1976, 1980] and Chew [1983]. Stochastic dominance in a more generalized nonlinear utility model, in contrast to expected (linear) utility theory is examined by Fishburn [1988].

Defining recursively $H_k(x) = {}_a\int^x H_{(k-1)}(x)dx$, $k \geq 1$, it is apparent that $H_1(x)$, which is the difference between the cumulative distributions $G(x)$ and $F(x)$, is non-negative on $[a,b]$, with $H_1(a) = H_1(b) = 0$. The original distribution thus dominates the transformed one by the First Degree Stochastic Dominance (FSD) criterion, given below in Section III.[3] The risky prospect under $f(x)$ will thus be preferred to $g(x)$ by all individuals for whom utility is increasing in x. The set of all DT(1) of $f(x)$ is a proper subset of transformations which make $f(x)$ preferred to $g(x)$ by FSD. The inclusion is proper since we have restricted the mass transferred to $[a, \mu-\epsilon]$ to be drawn from $(\mu-\epsilon, \mu+\epsilon]$.

In the context of this paper, the most important feature of a downward transfer is its effect on the central moments of the two distributions. As the decrease in the mean of $g(x)$ relative to $f(x)$ is obvious, we consider the effect of the downward transfer on the variance and the higher central moments of $g(x)$ relative to $f(x)$. Expressing both distributions in terms of deviations about their respective means, the difference between the k-th central moments of $g(x)$ and $f(x)$ is

$$\mu_g{}^k - \mu_f{}^k = {}_a\int^b x^k dH_1(x) \tag{1}$$

or, since $H_1(x)$ is increasing on $[a,-\epsilon]$ and is decreasing on $(-\epsilon, b]$,

$$\mu_g{}^k - \mu_f{}^k = x_1{}^k {}_a\int^{-\epsilon} dH_1(x) + x_2{}^k {}_{-\epsilon}\int^{\epsilon} dH_1(x) + x_3{}^k {}_{\epsilon}\int^{b} dH_1(x) \tag{2}$$

3 This criterion was originally discovered by Lehmann [1955] and introduced into the economic literature by Quirk and Saposnik [1962].

where $x_1 < -\epsilon < x_2 < \epsilon$, $\epsilon > 0$. The third integral term on the RHS is zero, since $H_0(x) = 0$ for $x > \epsilon$. Hence

$$\mu_g^k - \mu_f^k = (x_1^k - x_2^k) \int_a^{-\epsilon} dH_1(x) \tag{3}$$

As the integral in (3) is non-negative and $|x_1| \geq |x_2|$, $(-1)^i \mu_g^i \geq (-1)^i \mu_f^i$, for $i \geq k$, the even central moments of $g(x)$ are greater and the odd central moments less than those of $f(x)$.

A DT(1) corresponds quite closely to what non-economists often mean by an increase in risk. Economists and operations researchers, however, have generally preferred to restrict comparisons of risky alternatives to situations in which the means or other measures of location of the distributions are equal, so that preference for one over the other is due to variance and other characteristics of the distributions. One such characterization is the well-known "mean preserving spread" (MPS) of Rothschild and Stiglitz [1970]. The MPS is a special case of Second Degree Stochastic Dominance (SSD), which does not require equality of means.[4]

The addition of the restriction that the means be equal provides a definition of a downward transfer of the second degree, or DT(2), as a transfer of some of the mass of the distribution from an ϵ-neighborhood of the mean downward to $[a, -\epsilon]$, subject to $\mu_g^1 = \mu_f^1$. As was the case for a DT(1), the

4 SSD was discovered by Hardy, Littlewood, and Polya [1934] and was introduced to economists, apparently independently of each other and of earlier work, by Fishburn [1964], Hadar and Russell [1969], and Hanoch and Levy [1969].

probability of worst-case outcomes (i.e. $x \leq -\epsilon$) is increased. The set of DT(2) is a proper subset of MPS transformations, or of those transformations which make $f(x)$ preferred to $g(x)$ by SSD. Once again, the inclusion is proper, since we have constrained $H_0(x) = 0$ on $(\epsilon . b]$.

The downward transfer concept readily generalizes to higher orders, as we successively restrict the variance and increasingly higher central moments of the distributions to be equal:

Definition 1: A downward transfer of the k-th degree, DT(k), is a transformation of $f(x)$ into $g(x)$ such that $H_0(x) = g(x) - f(x) \geq 0$, for $x \leq \epsilon$, with strict inequality for at least one x in $[a . -\epsilon]$ and $H_0(x) = 0$ for $x > \epsilon$, subject to $\mu_g{}^i = \mu_f{}^i$, where $\mu_f{}^i$ and $\mu_g{}^i$ are the ith central moments of $f(x)$ and $g(x)$ and $i = 1 \ldots (k-1)$.

There is thus an increased probability of worst-case outcomes under $g(x)$, but shifts in mass on $(\epsilon . -\epsilon]$ leave the means and second through (k-1)st central moments unchanged.

The effects of downward transfers are less easily visualized as their order increases. An example of a DT(3), however, is provided by the "mean-variance-preserving transformation (MVPT) of Menezes, Geiss, and Tressler [1980]. Consider the two point distribution $g(x)$ in which Prob $(x = 0) = .25$ and Prob $(x = 2) = .75$. It differs from the two point distribution $f(x)$ given by Prob $(x = 1) = .75$ and Prob $(x = 3) = .25$ by a DT(3), since both distributions have a mean of 1.5 and a variance of .75, but the former one has a higher probability of the worst-case outcome of zero. Menezes, Geiss, and Tressler characterize $g(x)$ as having "increased downside risk" and show that $f(x)$ would be preferred for all utilities with a positive third derivative.

Higher order downward transfers may appear unusual, but in fact are easily produced using a linear programming approach. Let the points of increase x_i, $i = 1, ..n$ of the discrete distributions $f(x)$ and $g(x)$ have probabilities $p_i \geq 0$ and let $a_i(j) = x_i^j$, $j = 0, 1, ...(k-1)$. For DT(k) of second order or higher, we assume that the x_i are expressed as deviations about the mean. Then $g(x)$ differs from $f(x)$ by a DT(k) if it is a solution to the linear program

$$\text{Maximize} \quad p_1 \;\geq\; 0 \text{ , Subject to}$$

$$\sum_i^n p_i a_i(0) = 1, \quad \sum_i^n p_i a_i(1) = 0, \quad \sum_i^n p_i a_i(2) = \mu_f^2$$

$$\ldots \sum_i^n p_i a_i(k-1) = \mu_f^{(k-1)} \tag{4}$$

As an example, Table 1 below shows the distribution $f(x)$ and a series of downward transfers ranging from a DT(1) through a DT(6) where $x_i = -3, -2, \ldots 3$, (i = 1..6), produced by maximizing the probability of -3, the worst case outcome, subject to the increasing series of constraints on the moments.

As is clear from the table, the higher degree DT's differ from the original distribution by less than do lower degree DT's, a consequence of the imposition of the increasing constraints placed on the maximization problem. A downward transfer, particularly a higher degree one, is a theoretical curiosum, and not in itself a useful characterization of increasing risk in any practical sense, since the progressive addition of the constraints on the moments reduces the likelihood that such a pair of risky alternatives would ever be observed. The downward transfer concept, moreover, is not as general a definition of increasing risk as are the MPS, MVPT, or various SD rules, since each DT is a proper subset of the corresponding SD rule.

Table 1

K-degree Downward Transfers

x:	-3	-2	-1	0	1	2	3
f(x)	0	1/8	0	3/4	0	1/8	0
f(x) + DT(1)	1	0	0	0	0	0	0
f(x) + DT(2)	1/2	0	0	0	0	0	1/2
f(x) + DT(3)	1/12	0	0	2/3	1/4	0	0
f(x) + DT(4)	1/18	0	0	8/9	0	0	1/18
f(x) + DT(5)	1/40	0	1/4	1/2	1/8	1/10	0
f(x) + DT(6)	1/48	0	5/16	1/3	5/16	0	1/48

Downward transfers are, however, a fundamental theoretical construct for isolating the effect of an increase in the probability of worst case outcomes on all the moments of the distributions being compared, by impounding the effects of successively higher moments. In this sense the DT is a natural generalization of the idea of associating greater risk with less preferred distributions when the means are equal, as is done with the MPS, or when the first two moments are equal, as in the case of the MVPT. The DT makes precise the notion that an increase in risk can be defined as a *ceteris paribus* increase in the probability of the least desired outcomes, where the *ceteris paribus* condition corresponds to the requirement that a successively larger sequence of the lower moments of the two distributions be equal. The usefulness of the DT concept derives from the fact that each DT is a proper subset of a corresponding and more general increase in risk like the MPS or MVPT. It therefore serves as minimum condition, in that it appears reasonable to require that individuals who avoid worst-case outcomes at the very least prefer $f(x)$ to $g(x)$, where the latter differs from the former by a DT of any degree. In the following section downward transfers are used to characterize the utility function, under the hypothesis that the Pratt-Arrow risk premium should not decrease with a downward transfer of any degree.

II. The Risk Premium and Higher Moments

Following Pratt [1964], we consider the local risk premium π such that an individual would be indifferent between the risk $\tilde{z}$ and the non-stochastic amount $E(\hat{z})-\pi$. We maintain all of the assumptions of Pratt, except that we shall not assume that moments of the third or higher orders are negligible. Specifically, we shall assume that $\hat{z}$ has zero mean and is small in relation to the individual's base wealth x , so that terms in π of second and higher order may be ignored. The argument of this section assumes that x is non-stochastic. This assumption may be relaxed, following Machina and Neilson [1987], without essentially changing our argument. Throughout this and the following section we assume that $E[U(x)]$ exists and is finite.

Indifference between the certain and uncertain prospect requires that

$$U(x-\pi) = E[U(x-\hat{z})] \tag{5}$$

Expanding both sides of (5) about x , taking expectations of the RHS, and dividing by $-U^{(1)}(x)$, yields

$$\pi = \sum_{2}^{n}\left[\frac{-U^{(i)}(x)}{U^{(1)}(x)}\right]\left(\frac{1}{i!}\right)\mu^{i} \tag{6}$$

where μ^i is the ith central moment of the distribution of $\hat{z}$, since $E(\hat{z})=0$ and x is non-stochastic.

If the moments of $\hat{z}$ beyond the second order are ignored, so that $n = 2$, then for a given level of x the risk premium is linear in the variance of $\hat{z}$. Under this assumption, as Pratt and Arrow have shown, the term $r(x) = -U^{(2)}(x)/U^{(1)}(x)$, which is positive for all concave utilities, provides a natural measure of the degree of absolute risk aversion.

The generally accepted hypothesis that $r(x)$ is non-increasing in x is sufficient, but not necessary for $U^{(3)}(x) > 0$, a condition used by Whitmore [1980] to derive a third degree stochastic rule (TSD). Although further restrictions on the utility function might appear to have little meaningful economic content, the DT concept developed in Section I provides a rationale for such restrictions.

If we identify an increase in risk with any downward transfer, rather than with variance alone, then under the plausible hypothesis that the risk premium π should at least not decrease with an increase in risk, we can restrict the set of utilities to completely monotone concave utilities. This is the sense of Theorem 1 below, which shows that the hypothesis that the risk premium is non-decreasing for all downward transfers is equivalent to assuming that the derivatives of the utility function alternate in sign.

<u>Theorem 1</u>: Let $\pi(x.\tilde{z})$ be non-decreasing for all DT(k), $k = 1.2..n$. Then the utility function must satisfy the condition that

$$\frac{U^{(i+1)}(x)}{U^{(i)}(x)} \leq 0 \tag{7}$$

where $i = 1....n$.

<u>Proof</u>: The sufficiency of $U^{(i+1)}(x)/U^{(i)}(x) \leq 0$ for $i = 1....n$ and $\pi_g \geq \pi_f$ when $g(x)$ differs from $f(x)$ by a DT(k) for any $k \geq 1$ follows directly from (6), since $U^{(1)} > 0$, and Definition 1 implies that $\mu_g{}^1 = \mu_f{}^1$, $\mu_g{}^2 = \mu_f{}^2$, . . . $\mu_g{}^{(k-1)} = \mu_f{}^{(k-1)}$ and $(-1)^i \mu_g{}^k \leq (-1)^i \mu_f{}^k, i \geq k$.

To show the necessity of complete monotonicity for all $i = 1 \ldots. n$, rather than simply that $r^{(1)}(x) = -U^{(2)}(x)/U^{(1)}(x) \leq 0$, we show that it is always possible to produce a utility function for which $r^{(1)}(x) \leq 0$ but $U^{(i-1)}(x)/U^{(i)}(x) \geq 0$ on some subinterval of $[a.b]$, for a given $i \geq 3$, so that when x is in this subinterval a DT(i) implies that $\pi_g < \pi_f$

Let $\phi(x) = -\ln U^{(1)}(x)$. Since there is no danger of ambiguity in notation, we denote the successive derivatives of $\phi(x)$ by $\phi_1, \ldots. \phi_n$. Hence $\phi_1 = -U^{(2)}(x)/U^{(1)} = r(x)$. Decreasing absolute risk aversion in the conventional Pratt-Arrow sense requires that $\phi_2 = r^{(1)}(x) < 0$. By definition, $U^{(1)}(x) = e^{-\phi(x)}$. Successive differentiation of $U^{(1)}(x)$ yields the higher derivatives of $U(x)$, each of which is the product of $e^{-\phi(x)}$ and a function of the derivatives of $\phi(x)$.

The conditions $\phi_1 > 0$ and $\phi_2 < 0$ require that $U^{(2)}(x) < 0$ and $U^{(3)}(x) > 0$. For $i \geq 4$, however, the $U^{(i)}(x)$ may be of indeterminate sign on $[a.b]$, since they involve terms in at least ϕ_{i-1} and possibly higher, which may be of arbitrary sign. As we may always choose a function $\phi(x)$, subject to $\phi_1 > 0$ and $\phi_2 < 0$, we can produce a concave, decreasingly absolute risk averse utility which is not completely monotone, thus proving Theorem 1.

An example may make the argument more concrete. Consider the utility for which $\phi(x) = [\beta(\gamma \sin x + \cos x)/(1+\gamma^2) - \alpha/\gamma]e^{-\gamma x}$. Despite its rather bizarre form, the corresponding utility is strictly concave and decreasingly absolute risk averse for all $x \geq 0$ if $\alpha > \beta > 0$ and $\beta < \gamma\alpha$. For appropriate α, β and γ , each $U^{(i)}(x)$, $i \geq 4$, will vary in sign on $[0, \infty)$.

III. K-degree Stochastic Dominance

Let an individual with the utility function $U(x)$ face the random prospects 1 and 2, whose densities we denote by $f(x)$ and $g(x)$ respectively. As before, we assume that the random variable is bounded on $[a, b]$. The Von-Neumann-Morgenstern expected utility of the first prospect is then

$$E[U(x_1)] = \int_a^b U(x)dF(x) \tag{8}$$

with a corresponding definition for the second prospect. The first prospect is thus preferred if

$$\begin{aligned} E[U(x_1)] - E[U(x_2)] &= \int_a^b U(x)[f(x) - g(x)]dx \\ &= -\int_a^b U(x)dH_1(x) > 0 \end{aligned} \tag{9}$$

Integrating by parts, and noting that $H_1(a) = H_1(b) = 0$,

$$E[U(x_1)] - E[U(x_2)] = \int_a^b H_1(x)U^{(1)}(x)dx > 0 \tag{10}$$

Since $U^{(1)}(x) > 0$, a sufficient condition for the first prospect to be preferred is that

$$H_1(x) \geq 0, \ \forall x \in [a, b] \tag{11}$$

with strict inequality for at least one $x \in [a.b]$. This condition, which is usually termed First Degree Stochastic Dominance, requires that the distribution $F(x)$ lie wholly below $G(x)$ at all points on $[a.b]$. For this reason FSD is of limited practical usefulness in ordering risky alternatives, since the distribution of any pair of non-identical prospects with equal means will intersect at least once. In addition, as is the case for all lower-degree SD rules, $H_1(x)$ must be evaluated at all points on $[a.b]$

A more slightly more powerful rule may be obtained from (10) by again integrating by parts, yielding

$$\left|U^{(1)}(x)H_2(x)\right|_a^b - \int_a^b U^{(2)}(x)H_2(x)dx \tag{12}$$

For risk averse ($U^{(1)}(x)>0, U^{(2)}(x)<0$) utilities, a sufficient condition for preference for the first prospect is that

$$H_2(x) \geq 0, \quad \forall x \in [a,b] \tag{13}$$

with strict inequality for at least one $x \in [a.b]$. This condition, or Second Degree Stochastic Dominance (SSD) reduces to Rothschild and Stiglitz' "integral conditions" for a mean-preserving spread when $\mu_f{}^1 = \mu_g{}^1$, since $H_2(b) = H_2(a) = 0$. The greater power of this rule, which orders a larger set of uncertain prospects, results from the additional restriction that $U^{(2)}(x)<0$. Analogously to FSD, it requires the evaluation of $H_2(x)$ for all $x \in [a.b]$.

In the preceding section it was shown in Theorem 1 that the hypothesis that the risk premium π should be non-decreasing for any downward transfer implies complete monotonicity. The additional restrictions on utility permit the derivation of SD rules up to any desired degree. Continued integration by parts of (12) yields the general k-th degree expression for the difference in the expected utilities, or $E[U(x_f)] - E[U(x_g)]$,

$$= -\sum_{i}^{k}(-1)^{i}U^{(i)}(b)H_{i+1}(b) - (-1)^{k-1}\int_{a}^{b}U^{(k+1)}(x)H_{k+1}(x)dx \tag{14}$$

The corresponding k-th degree SD rule based on (14) is thus that

$$H_{k+1}(x) \leq 0, \quad \forall x \in [a,b] \tag{15}$$

with strict inequality for at least one $x \in [a.b]$, and $H_i(b) \leq 0$, $i = 1,..k$.

A k-th degree SD rule is theoretically superior to lower degree SD rules in that it orders a larger set of distributions. At an operational level, moreover, it has an added advantage, since it reduces the choice between two alternatives to an evaluation of the $H_i(b)$, $i = 1,..k$ only at b, rather than over all $x \in [a.b]$. This follows from the fact that for finite expected utilities, the residual integral term in (14) tends to zero as k increases, requiring that only the sequence of $H_i(b)$ be non-negative (non-positive) for $E[U(x_1)] \geq (<=) E[U(x_2)]$. Since the $H_i(b)$ are linear in the means and central moments of the distributions of the two prospects, the problem of expected utility choice is reduced to a comparison of moments, provided that they exist.

References

Arrow, Kenneth: *Essays in the Theory of Risk Bearing*, Helsinki, Yrjö Jahnssonin Säätiö, 1965

Chew, Soo Hong: "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox", *Econometrica* **51** (July 1983) 1065-1094

Fishburn, Peter: *Decision and Value Theory*, New York, John Wiley & Sons, 1964

Fishburn, Peter: "Continua of Stochastic Dominance Relations for Bounded Probability Distributions", *Journal of Mathematical Economics* **3** (December 1976) 295-311

Fishburn, Peter: "Continua of Stochastic Dominance Relations for Unbounded Probability Distributions", *Journal of Mathematical Economics* **7** (December 1980) 271-285

Fishburn, Peter: "Stochastic Dominance and Moments of Distributions", *Mathematics of Operations Research*, **5** (February 1980), 94-100

Fishburn, Peter: *Nonlinear Preference and Utility Theory*, Baltimore, The Johns Hopkins University Press, 1988

Hadar, Josef, and William Russell: "Rules for Ordering Uncertain Prospects", *American Economic Review*, **69** (March 1969), 25-34

Hanoch, Giora and Haim Levy: "The Efficiency Analysis of Choices Involving Risk", *Review of Economic Studies*, **36** (July 1969), 335-346

Hardy, G.H., J.E. Littlewood, and G. Polya: *Inequalities*, Cambridge, Cambridge University Press 1934

Kihlstrom, Richard E. and Leonard J. Mirman: "Constant, Increasing and Decreasing Risk Aversion with Many Commodities", *Review of Economic Studies* **48**, (April 1981), 271-280

Lehmann, E.L.: "Ordered Families of Distributions", *Annals of Mathematical Statistics*, **26** (September 1955), 399-419

Machina, Mark J.: "A Stronger Characterization of Declining Risk Aversion", *Econometrica* **50** (July 1982) 1069-1079

Machina, Mark J. and William S. Neilson: "The Ross Characterization of Risk Aversion: Strengthening and Extension", *Econometrica* **55** (September 1987) 1139-1149

Menezes, Carmen, Charles C. Geiss, and John H. Tressler: "Increasing Downside Risk", *American Economic Review* **70** (December 1980) 921-932

Pratt, John: "Risk Aversion in the Small and in the Large", *Econometrica* **32** (January-April 1964) 122-136

Pratt, John and Richard J. Zeckhauser: "Proper Risk Aversion", *Econometrica* **55** (January 1987) 143-154

Quirk, James P. and Rubin Saposnik: "Admissibility and Measurable Utility Functions", *Review of Economic Studies* **29** (February 1962) 140-146

Rothschild, Michael and Joseph Stiglitz: "Increasing Risk: A Definition", *Journal of Economic Theory*, **2** (September 1970) 225-243

Ross, Stephen A.: "Some Stronger Measures of Risk Aversion in the Small and the Large with Applications", *Econometrica* **49** (May 1981) 621-638

Whitmore, G.A.: "Third Degree Stochastic Dominance", *American Economic Review* **60** (June 1970) 457-459